JOIV : Int. J. Inform. Visualization, 9(2) - March 2025 808-816

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Crypto Forecast: Integrating Web Scraping and Data Analysis for Cryptocurrency Price Prediction

Krutika Gadge [a], Shreyash Daduria [a], Abhishek Sarodaya [a], Pradnya Borkar [a,*], Sagarkumar S. Badhiye [a], Pratik K Agrawal [a]

[a] *Symbiosis Institute of Technology, Nagpur Campus, Symbiosis International (Deemed University), Pune, India,*
Corresponding author: [*]pradnyaborkar2@gmail.com

***Abstract*—Accurately predicting cryptocurrency prices is still a difficult task because of the extremely volatile nature of the market. This study introduces a new methodology combining web scraping, data analysis, and machine learning to further improve prediction accuracy. A live cryptocurrency monitors gathers data from various sources such as trading volumes, price volatility, and sentiment in market to create a rich data set. Feature engineering is used to convert raw data into useful inputs for machine learning algorithms to further enhance prediction functions. Utilizing Python libraries including Beautiful Soup, Pandas, Scikit-learn, and deep learning libraries, the correct predictive model is designed and strictly tested for precision, performance, data quality, usability, scalability, and cost. The proposed hybrid model is a combination of traditional statistical methods with deep learning models to overcome the constraints of conventional forecasting methodologies. The output reflects the performance of the model in identifying the trends in the market and rendering data-driven insights to traders and investors. Future studies can employ different data sources, including social media sentiment analysis, financial news articles, and web-based cryptocurrency forums, to enhance predictability. Further advancement in time series forecasting through deep learning models, including transformer models, may also enhance the precision of long-term forecasting. A deeper insight into how external forces, including government intervention, macroeconomic trends, and emerging blockchain technologies, would complement our understanding of cryptocurrency market dynamics. This study contributes to complementing predictive analytics in financial markets by providing useful insights to investors, researchers, and policymakers.**

***Keywords*—Cryptocurrency; web scraping; crypto forecast.**

Manuscript received 30 May 2024; revised 6 Jun. 2024; accepted 28 Jul. 2024. Date of publication 31 Mar. 2025.

I. Introduction

Cryptocurrencies are digital or virtual currencies that use cryptography for secure financial transactions. Unlike traditional currencies controlled by intermediaries like banks or governments, cryptocurrencies operate on blockchain-based decentralized networks, eliminating the need for intermediaries. Bitcoin, introduced in 2009 by an anonymous individual or group using the pseudonym Satoshi Nakamoto, was the first and most well-known cryptocurrency. Since then, thousands of other cryptocurrencies have been developed. The emergence of these digital assets has had a profound impact on the financial landscape. It has disrupted traditional financial systems because of its several advantages, which include decentralization, Transparency, Financial Inclusion, Investment Opportunities, Trading, pseudo-anonymity, and Global accessibility. The total market capitalization of cryptocurrency has grown significantly, reaching over 2 trillion at its peak in 2021 [1].

Their potential for high returns and decentralized transactions have made this digital asset a disruptive force in the financial landscape; its growing impact presents several challenges for analyzing cryptocurrency markets and making it a challenging task. Challenges include extreme volatility, susceptibility to speculations, Lack of historical data, Regulatory uncertainty, Market fragmentation and manipulation, and technological risks, making it difficult to analyze and predict trends accurately. These challenges add complexity to understanding and predicting market dynamics. Therefore, data-driven approaches are crucial for gaining meaningful insights. Collecting and analyzing datasets, Handling volatility, Risk management, objective analysis, Identifying patterns and trends, and Adaptability and flexibility.

The aim is to collect cryptocurrency data and leverage the data analysis techniques to understand cryptocurrency market trends better. This would help us to unlock valuable insight from the cryptocurrency market. One of the key goals is data collection, which aims to collect comprehensive data such as price history, market capitalization, and trading volume from various cryptocurrencies from a reputable source such as CoinMarketCap. For this project, we have chosen a time frame for data collection from April 2013 to April 2024. The data frame selected for data scraping spans multiple years, allowing us to capture significant historical perspectives and account for various market conditions, trends, and volatility witnessed, allowing us to analyze trends and identify potential patterns.

Once the data is collected, it intends to utilize various time series analysis techniques to gain meaningful insights into the cryptocurrency price movements and forecast future trends, including Trend analysis, Volatility analysis, Time series analysis, and sentiment analysis. Each technique will be applied based on its suitability to the specific dataset, and by incorporating these data-driven techniques, the project aims to gain a deeper understanding of cryptocurrency market trends, identify patterns and potential opportunities, and provide insights for investors, traders, and policymakers for more informed decision-making.

Cryptocurrencies have upended established ideas about money and payment methods in the financial industry, bringing forth a new paradigm. While their impact continues to evolve and add complexity to understanding, using data-driven approaches would help us gain insights. It's evident that cryptocurrencies are changing the financial environment [2] and spurring innovation across a range of industries, even though their influence is still developing.

Cryptocurrency price prediction [3] anticipates future values of cryptocurrencies including Bitcoin, Ethereum, and other digital currencies. Accurate price prediction is essential for the cryptocurrency market, which is very volatile; it helps investors make informed decisions, manage risks, and avoid losses. Nevertheless, predicting cryptocurrency prices is very difficult because the market is complex, random, and reacts to many factors, such as market sentiment, changes in regulations and emerging technologies.

Historically, the forecasting of cryptocurrency prices [4] is based on technical analysis, which includes examining historical prices and volume data to find patterns and trends, and fundamental analysis, which evaluates the technology, adoption rates, and market sentiment, respectively. On the other hand, these methods are widely used, but they are still falling short of capturing the complex character of the cryptocurrency market, which contributes to inappropriate predictions.

In recent times, researchers have tried different and more sophisticated methods, including machine learning techniques, sentiment analysis, and web scraping in their models to improve the accuracy of cryptocurrency forecasting models. These methods harness a large amount of data from different sources, such as social media networks, online forums, and news channels, to create useful information about price changes.

Sentiment analysis is a tool for assessing opinions and emotions expressed in data sources such as news, social media, and forums. It employs natural language processing (NLP) to interpret text-based data and calculate sentiment polarity (positive, negative, or neutral). Techniques such as lexicon-based (using predefined dictionaries) or machine learning-based (learning from labeled datasets) might be used.

Research shows that sentiment has a substantial influence on asset prices and trading behavior, so positive sentiment sparks the buying force and rising prices, while negative sentiment leads to selling and falling prices. Nonetheless, they involve context-sensitive language, ironic remarks, inherent biases, and complexities beyond the scope of sentiment analysis [5], [6].

Machine learning techniques are increasingly acknowledged as powerful instruments in financial modeling, most notably in predicting cryptocurrency prices. Unlike traditional methods, machine learning can separate complex, non-linear relations within large datasets like random forests, gradient boosting, and LSTM networks.

Machine learning models beat conventional methods by adjusting to changes in market circumstances. However, problems remain, such as the need for many large data sets and dealing with market volatility. Methods like regularization and cross-validation can be used to tackle problems like overfitting [7].

Although these barriers exist, the attempt to get more precise predictive models [8], [9] for cryptocurrency price fluctuations is very alive. The infusion of machine learning and financial analysis creates promising channels for optimizing predictions and handling the complicated nuances in crypto markets. The small number of stocks in some industries limits the analysis. Outlier analysis should be used in future studies to choose defensive Sharia-compliant stocks that perform well, as discussed in [10]. To help the marketing department's decision-making, the development can concentrate on certain users such as data analysts, business intelligence analysts, or other positions inside an organization [11].

Web scraping tools have evolved into a vital component in collecting cryptocurrency data, allowing analysts to extract information from various online resources. Using data extraction from platforms, including cryptocurrency exchanges and social networks, scientists gain real-time and past data, which are paramount to market dynamics analysis. The speed of web scraping is crucial for such a fast-paced cryptocurrency environment, which helps make smart trading decisions. This way, data can be drawn from different sources, such as trading volumes and market capitalization.

On the one hand, the ethical aspect is a major problem. Any scraping without an appropriate schedule can be considered against the website terms or may generate an unnecessary number of requests. Respecting the defined responsible practices, such as observing robots.txt files and abiding by website policies, is the key to cooperating with data providers.

Web scraping provides real-time data and allows for the collection of historical data for model training. This is essential for the development of machine learning models as it helps with, for instance, social media sentiment analysis integrating with database enrichment, which increases predictive capabilities.

Yet, challenges persist. Disparities in website structures [12] often result in data quality issues, and legal complexities

are characterized by complying with privacy and intellectual property laws. Knowledge of coding languages like Python and tools like Requests and Beautiful Soup are indispensable for successful scraping. Furthermore, data cleaning and preprocessing allow for the accuracy of scraped data.

Several studies [13] have explored incorporating web scraping, data analysis, and sentiment analysis into cryptocurrency price forecasting, taking advantage of the strengths of all the approaches. This integrated approach involves examining complete market dynamics and opinions, leading to more precise and reliable forecasting models.

Recent research suggested that methods integrating web scraping for data extraction from multiple sources, sentiment analysis using NLP techniques, and machine learning techniques like random forests and neural networks for price prediction were better. The study has proved to be successful, overtaking traditional methods consisting of technical only or fundamental analysis.

Innovative methods, such as ensemble models that combine multiple algorithms and data sources, have made prediction models even more accurate and reliable. Future research could encompass more advanced NLP techniques, alternative data sources such as news events or social media influencer activity, and ensemble models developed by combining different data types and modeling approaches.

With the increase in computational resources and data availability, researchers could employ more sophisticated Machine Learning models, such as deep learning architecture, web scraping and data analysis techniques, for cryptocurrency price prediction models. Such integrated procedures may result in better forecasting accuracy and the development of real-time systems.

Cryptocurrency price prediction through web scraping, data analysis, and sentiment analysis determines accuracy, while conventional methods do not. Current machine-learning algorithms, including neural networks, random forests, and support vector machines, can perform better on price modeling of cryptocurrencies due to analyses of large volumes of past data. Web scraping receives current and historical data from various online platforms using different tools, heightening forecasting accuracy. Complementary architectures of the scraped data with sentiment processing have demonstrated superior abilities, and ensemble models have displayed better predictions. Besides, future studies must focus on sophisticated natural language processing for sentiment analysis and alternate data sources such as news events [14], [15]. By applying deep learning models alongside web scraping and data analytics, real-time forecast systems and prediction accuracy will probably be enhanced [16]. This multidisciplinary approach is expected to result in substantial progress in cryptocurrency prediction models and, hence, will aid investors, traders, and other actors in this market.

II. Materials and Methods

A. Data Acquisition: Web Scraping with Coinmarketcap

Web scraping can be defined as the automatic extraction of data from websites, which is quite interesting in the areas of financial data analysis and cryptocurrency price prediction. By applying web scraping methods, scientists and investigators may obtain data in real-time or historical form from different online sources, including cryptocurrency platform sites, news platforms, and social media platforms.

Yet, the process of web scraping should always be accompanied by some strong awareness of ethical and legal ramifications. The sites usually have jurisdictions that stipulate the rules of collecting the data and how it will be used. Regarding the CoinMarketCap case, which often takes place among cryptocurrency data aggregators, their terms of service do not allow the automated use of web scrapers without prior consent. Python is popular for web scraping, it offers several libraries for the purpose of web scraping:

a. Beautiful Soup: XML and HTML files parsing libraries that are useful in extracting data from web-pages based on tags, attributes and classes has reduced the, complexity of the data extraction exercise,
b. Scrapy: A robust framework with the capability of renders, tweak selective HTML elements, processing, and extracting information,
c. Pandas: One of the reasons for its preference over other languages is that it is mainly utilized for data cleaning and analysis purposes after scraping the data.
d. The crypto data which will be used for this research project will be collected from CoinMarketCap, a well-known cryptocurrency data aggregator. For this to happen, we have used some Python-based web scraping libraries like Beautiful Soup and Scrapy, which are relatively easy to work with and provide a rich set of tools for extracting data from web pages.

The data scraping process will involve the following steps:

1) Developing the protocol for locating target coin addresses on CoinMarketCap for the given data numbers-cryptocurrency prices, trading volumes, and market capitalizations [17].

2) Pull the nest data set from the HTML structure of the target web pages by exploiting the offered parsing techniques by the selected web scrapping packages.

3) Anti-algorithms and data protection systems in place to ensure the quality and completeness of the data are addressed, including handling network errorsork, page changes, or missing data.

4) Powerful scraper used to keep and format the scraped data in a structured way e.g. CSV, JSON, for further processing and analysis as in Fig. 1.

Fig. 1 Data scraping process

Data collection will be followed by data pre-processing steps which will entail missing value handling, data or formatting. We will do this to avoid the risk of experiencing data errors and being able to handle feature engineering, prediction modelling, etc.

The implementation of the grants includes data from CoinMarketCap to both other aspects, ethical, and legal, acquainting with the website's terms of service, utilizing the

proper tools, and following the steps, developers would efficiently extract and process data for further analysis as in Table I.

B. Exploratory Data Analysis (EDA)

After extracting CoinMarketCap crypto data, the next phase is to perform an in-depth exploratory data analysis (EDA) to enable deeper data scrutiny and reveal hidden patterns within our dataset. The scraped data covers features of cryptocurrencies, including name, symbol, price, market cap, and volume. Bar charts are used to track the cryptocurrencies' price, market cap, and trading volume over time, giving a glimpse of the performance and volatility of the most prominent cryptocurrency.

Histograms and bar graphs could show the distribution of key features like price and market cap. Histograms and bar graphs depict the distribution of these variables, visualizing the shape of the data using skewness and modality as well. Bar graphs are suitable for detecting outliers and understanding the extent of the spread ranging from the minimum to maximum values.

TABLE I
SCRAPED DATA

	DATE	NAME	SYMBOL	MARKET CAP	PRICE	CIRCULATING SUPPLY	VOLUME (24HR)	% 1H	% 24H	% 7D
0	/historical/20130428	Bitcoin	BTC	$1,488,566,971.96	$134.21	11,091,325	None	0.64	…	…
1	/historical/20130428	Litecoin	LTC	$74,637,021.57	$ 4.3484	17,164,230	None	0.80	…	…
2	/historical/20130428	Peercoin	PPC	$7,250,186.65	$ 0.3865	18,757,230	None	-0.93	…	…
…	…	…	…	…	…	…	…	…	…	…
5714	/historical/20240407	Dogecoin	DOGE	$28,630,384,308.50	$ 0.1991	143,792,996,384	$2,643,420,653.78	1.28	7.08	-9.52
5715	/historical/20240407	Cardano	ADA	$20,985,760,903.89	$ 0.5895	35,600,844,566	$307,767,311.07	0.78	0.88	-9.37
5716	/historical/20240407	Toncoin	TON	$18,637,549,799.82	$ 5.3701	3,470,633,063	$106,856,795.84	-0.06	-1.99	-0.83

1) Missing Data: In addition, missing data were also detected and handled accordingly. A function that skipped dates for which data were not available was also added to the code. Missing values in the "Volume (24hr)" column are found using the df.isnull().sum() method [18]. To ensure the "Volume (24hr)" column is complete, the K-Nearest Neighbors (KNN) imputation approach was chosen. KNN imputation is a more accurate method than the mean and median imputations and is capable of predicting and filling in the missing values [19], as shown in Table II and Table III.

TABLE II
BEFORE MISSING VALUE IMPUTATION

	Date	Name	Symbol	Market Cap	Price	Circulating Supply	Volume (24hr)	% 1h	% 24h	% 7d
0	2013-04-28	Bitcoin	BTC	1488566972	134.21	11091325	NaN	0.64	0	0
1	2013-04-28	Litecoin	LTC	74637021.57	4.35	17164230	NaN	0.8	0	0
2	2013-04-28	Peercoin	PPC	7250186.65	0.39	18757362	NaN	0.01	0	0
3	2013-04-28	Namecoin	NMC	5995997.19	1.11	5415300	NaN	0.01	0	0
4	2013-04-28	Terracoin	TRC	1503099.4	0.65	2323570	NaN	0.61	0	0
...		-	...	...	...	...	...	...	...	
5712	2024-04-07	USDC	USDC	32936774705	1	32939363285	4713439433	0	-0.01	-0.01
5713	2024-04-07	XRP	XRP	32746512775	0.59	55051549471	924956308.5	0.38	0.25	-5.47
5714	2024-04-07	Dogecoin	DOGE	28630384309	0.2	1.43793E+11	2643420654	1.28	7.08	-9.52
5715	2024-04-07	Cardano	ADA	20985760904	0.59	35600844566	307767311.1	0.78	0.88	-9.37
5716	2024-04-07	Toncoin	TON	18637549800	5.37	3470633063	106856795.8	0.01	-1.99	-0.83

TABLE III
AFTER MISSING VALUE IMPUTATION

	Date	Name	Symbol	Market Cap	Price	Circulating Supply	Volume (24hr)	% 1h	% 24h	% 7d
0	2013-04-28	Bitcoin	BTC	1488566972	134.21	11091325	94460947.96	0.64	0	0
1	2013-04-28	Litecoin	LTC	74637021.57	4.35	17164230	108629405.9	0.8	0	0
2	2013-04-28	Peercoin	PPC	7250186.65	0.39	18757362	5939582.36	0.01	0	0
3	2013-04-28	Namecoin	NMC	5995997.19	1.11	5415300	5939582.36	0.01	0	0
4	2013-04-28	Terracoin	TRC	1503099.4	0.65	2323570	90179.32	0.61	0	0
...		...	...	...	...	...	...	...	...	
5712	2024-04-07	USDC	USDC	32936774705	1	32939363285	4713439433	0	-0.01	-0.01
5713	2024-04-07	XRP	XRP	32746512775	0.59	55051549471	924956308.5	0.38	0.25	-5.47
5714	2024-04-07	Dogecoin	DOGE	28630384309	0.2	1.43793E+11	2643420654	1.28	7.08	-9.52
5715	2024-04-07	Cardano	ADA	20985760904	0.59	35600844566	307767311.1	0.78	0.88	-9.37
5716	2024-04-07	Toncoin	TON	18637549800	5.37	3470633063	106856795.8	0.01	-1.99	-0.83

2) Initial Insights: Bitcoin, the most well-known cryptocurrency, displayed noticeable volatility in its price, market cap, and trading volume over time, illustrating the rising and falling character of the cryptocurrency market. Some cryptocurrencies were found to have strong positive correlations between their market cap and their prices, which implies that the market cap plays a vital role in affecting or reflecting the price. The market datasets, which contain a distribution of price and market caps among different cryptocurrencies, also identified

potential outliers, such as cryptocurrencies with exceptionally high or low values that needed more investigation into their underlying factors or market dynamics. Volume in trading of certain cryptocurrencies can be a factor in their movements and sentiments in the market, presenting more opportunities for investigation.

C. Methodology

In this initiative, we implemented a combined approach of web scraping, data analysis, and predictive modeling to improve our understanding and forecasting of cryptocurrency prices. After extracting CoinMarketCap crypto data, the next phase is to perform an in-depth exploratory data analysis (EDA) to enable deeper data scrutiny and reveal hidden patterns within our dataset.

1) Web Scraping: The initial phase of the methodology was to collect the required data via web scraping. Utilizing Python APIs such as Requests and BeautifulSoup, along with data from CoinMarketCap to access historic cryptocurrency data. This process included determining the target URLs and HTML structures to scrape discrete values, such as prices, trading volumes, market capitalizations, and other key metrics. For ethical and responsible web scraping, the researchers considered the rules specified in CoinMarketCap's terms of service, respected the robots.txt file, and implemented the rate-limiting mechanisms. The data was extracted in a structured format and then processed and analyzed.

2) Data Preprocessing and Exploratory Analysis: After the data collection phase, in-depth data preprocessing and exploratory data analysis (EDA) were conducted [20]. The paper also discussed the problem of missing values, deleting unnecessary data, and type conversion to make clean data prior to further analysis.

The EDA encountered exploratory data analysis, comprising statistical calculations, data visualization techniques, and analysis of the quantitative relationships among the variables. The correlation analysis technique, the trend identification method, and the statistical volatility estimation method could unveil the inner dynamics of the crypto market.

3) Predictive Modelling: The final stage was centered around principally building and assessing the effectiveness of predictive models for cryptocurrency price predictions. Predictive Modelling: Once historical data is handed, predicting trends for the present and the future become much simpler. Specifically, a Long Short-Term Memory (LSTM) model is learned using the Bitcoin spot price data, which is available at https://coinmarketcap.com/currencies/bitcoinspot/. This model provides forecasts for current and future values, which are depicted against original data points to hint at price movements. Admittedly, this study gives us a good insight into the market's future as we predict based on the available limiting factors, which may change the precision rate as we go. The framework consists of pre-processing and normalizing data as a prelude to the LSTM model development with the training set. Forecasts are made with the help of those models, saving time for doing graphs to exhibit future price trends visually. The predictive approach encourages educated decision-making but with delegated responsibility out to chance, given the market unpredictability of the cryptocurrency world.

4) Graphical User Interface

The application developed to analyze cryptocurrency data had a graphical user interface (GUI) developed using the Tkinter library. The main window for the application contains many frames carrying different components. From top to bottom, the main window consists of a heading label, an output frame, and a button frame. The output frame displays several outputs, from the scraped data to visualization and analysis results. The heading label will have the title of the application: "Cryptocurrency Data Scraper. The button frame is filled as much as possible so that the user can interact with it to perform a different operation in the application. The user can initiate the web scraping process by clicking the "Start Scraping" button. The user then displays the scraped data in a text area after scraping data by clicking the "Display Data" button. The results of action/step are displayed in a text area together with a vertical scroll bar, which is implemented with a Text widget from Tkinter. The application provides several buttons for visualizing different aspects of the data. Clicking on the "Time Series Graph" button, entering the name of any cryptocurrency, and hitting the button will show the time series plot with the prices, market cap, and 24-hour trading volume over time for that coin. It uses the Matplotlib library to create these plots and widgets from the module matplotlib. backends. backend_tkagg to embed the created plots into the GUI. The other buttons of the button frame allow a user to view and analyze all sorts of information, including correlation matrixes, growth rates, volatility indexes, comparative data, and predictive analysis.

In some of the analyses—volatility index and predictive analysis—the app asks for the name of the cryptocurrency from the user using a simple dialog. ask string () from the Tkinter library. The 'pack()' method from Tkinter controls the layout of the GUI by placing widgets in either a vertical or horizontal form. The application is styled further, looking much more attractive with some visual stylings, such as the setting of the background color of frames and various widgets, font style, and color of labels and buttons. Generally, the interface was designed to interact with the user and with the application to analyze data on cryptocurrencies, enabling the most basic functionalities to be performed using web scraping. Plots represent many data aspects by combining Tkinter widgets with Matplotlib visualizations.

III. Result and Discussion

A. Time Series Analysis

Time series analysis takes one of the lead positions in analyzing cryptocurrency price dynamics throughout time. However, the study draws on the example of Bitcoin, the most popular cryptocurrency, analyzing its USD value, capitalization, and transaction volume charts. We needed to discover info about Bitcoin's past management and forecast its future performance using the data. The code used the module Matplotlib to plot the time series data, converted the 'Date' column to DateTime format, and restricted the information that involved Bitcoins. This customized plot format helped with readability, visually showing how the sub-plots line graphs of Bitcoin's price, market cap, and 24-hour'

volume moved against time. Date formatting contributed to the clearer unrolling of time frames [21], as shown in Fig. 2. And Fig. 3.

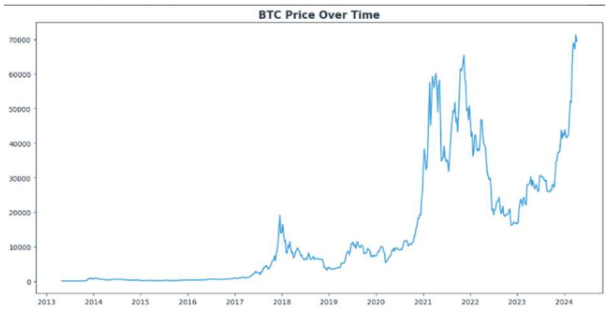

Fig. 2 BTC Price Over Time

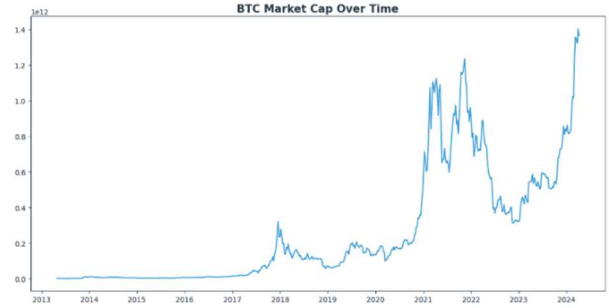

Fig. 3 BTC Market Cap Over Time

B. Additional Analysis

Moreover, different analytical methods, alongside time series analysis, were employed to understand the market dynamics and behavior of cryptocurrencies.

TABLE IV
CORRELATION ANALYSIS

	Market Cap	Price	Circulating Supply	Volume (24hr)	% 1h	% 24h	% 7d
Market Cap	1	0.95	0	0.4	-0.03	0	-0.01
Price	0.95	1	0	0.34	-0.02	0	-0.01
Circulating Supply	0	0	1	0	0.02	0	0.01
Volume (24hr)	0.4	0.34	0	1	-0.06	-0.01	-0.01
% 1h	-0.03	-0.02	0.02	-0.06	1	0.21	0.06
% 24h	0	0	0	-0.01	0.21	1	0.1
% 7d	-0.01	-0.01	0.01	-0.01	0.06	0.1	1

C. Correlation Analysis

The computation of the correlation matrix was implemented to define the relationships among key financial indicators, such as market capitalization, price, circulating supply, trading volume, and various percentage changes. This study showed that market capitalization and price are strongly related to the higher cost of cryptocurrencies with larger market caps. Market capitalization and trading volume also exhibited a moderate positive correlation, indicating that larger cryptocurrencies generally have higher trading activity. On the other hand, the short-term percentage changes in price (1-hour, 24-hour, 7-day) did not reveal strong correlations with the other variables, implying that some variables beyond the ones being considered may affect the short-term price movements, as shown in Table IV.

The growth rates for price and market capitalization were calculated to study the changes in cryptocurrency performance over time. By sorting the data by cryptocurrency name and date, they were able to identify the percentage changes of these main metrics for each cryptocurrency across the selected time frame. This analysis provided insights into which cryptocurrencies experienced significant growth or decline in price and market capitalization, revealing the varying trends and performance of different digital assets, as shown in Table V.

TABLE V
TREND ANALYSIS

Name	Price	Market Cap
Aphroditecoin	0	0
Augur	140.7727389	140.8016501
Auroracoin	-96.54217854	-96.50241377
Avalanche	-49 74458995077552	-22.0083333
BNB	17.18797023	5.058300365
Banx	56.24324099	93.21115828
Binance Coin	4.66361E+15	5528.298958
Binance USD	0.299850075	-15.5387676
BitConnect	17 644011603812686	20.05916392
BitShares	3553 6585365853653	46435986774631765
BitShares PTS	-38.06151049	3.380269991
Bitcoin	51582.10267	91585.2085
Bitcoin Cash	429 65195323121543	502.1328663
Bitcoin Gold	-37601897543553534	-3.75041E+15
Bitcoin SV	64.03307761	74.17489213
BlackCoin	-3569651741293532	-35.59397089
Bullion	235.9659425	302.7666999
Bytecoin	3142.24237	3697.288734
CHNCoin	-73.27122153	-72.36015296
Cardano	347.2685888	514.0890095
Chainlink	149.5908347	318.6496712
Counterparty	-54 51373133683368	-54 54927401341384
Crypto.com Coin	25.02099076	50.24443034
Dash	61520671471260335	115230.2849
Devcoin	-8483900643974242	-81 31001027816201

D. Volatility Analysis

To reflect the changing dynamics in the cryptocurrency price, the trios take the squared root of the squared average percent return on every digital currency. This also helps them know the implied volatility, which means how much the prices could move. This volatility analysis helps assess the stability and risk associated with different digital assets, enabling investors and traders to make more informed decisions regarding their portfolio allocations and risk management strategies, as depicted in Table VI.

TABLE VI
VOLATILITY ANALYSIS

Name	Volatility Index
Aphroditecoin	NaN
Augur	0.16
Auroracoin	0.41
Avalanche	0.17
BNB	0.08
…	…
Waves	7.30
WorldCoin	2.25
XCurrency	NaN
XRP	0.34
Yacoin	NaN

E. Comparative Analysis

To better compare cryptocurrencies, the most recent data for each digital asset, including the price, market capitalization, and trading volume, is used. Comparative

analysis makes it possible to estimate the efficiency and market activities of different cryptocurrencies; as a result, traders and investors have more options in their decision-making by paying attention to the general state of the crypto market, as in Table VII.

TABLE VI
COMPARATIVE ANALYSIS

Name	Price	Market Cap	Volume(24hr)
Aphroditecoin	3.85	86892531.75	45328.90
Augur	17.00	187021274.57	874532.06
Auroracoin	0.84	9051803.63	15410.31
Avalanche	54.11	20422170640.27	343450748.28
BNB	584.10	87343737468.08	1374676397.27
...	...	...	...
Waves	5.35	534606266.02	4447074.00
WorldCoin	0.52	19546847.28	123416.81
XCurrency	1.81	9965987.76	627465.00
XRP	0.59	32746512775.46	924956308.50
Yacoin	0.03	108880.76	5535361.18

F. Predictive Analysis

To accurately forecast future bitcoin prices, the researchers have used machine learning and created a Long Short-Term Memory (LSTM) model. Due to training the LSTM model on historical prices, researchers obtained forecasts that were not just for the original time frame but also for other periods in history. This predictive analysis provides insight into the direction in which future prices may go. However, I would like to point out here that this can only be a prediction as long as there is volatility and some limitations in accurately forecasting the highly volatile cryptocurrency market [22], as shown in Fig. 4.

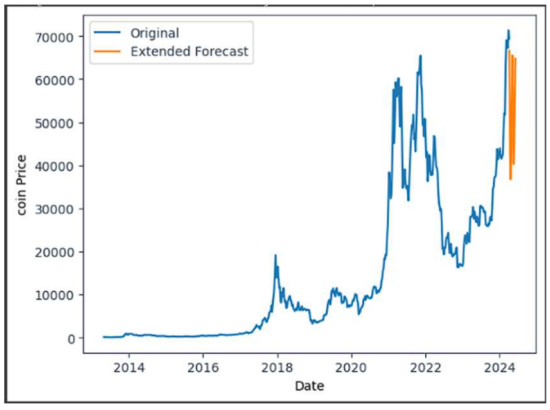

Fig. 4 Predictive Analysis

G. Investigation

The investigation comprehensively responded to the topic, examining the cryptocurrency market rules and flexible price determinants.

1) Key Findings: In the framework of the data gathering, the web scraping process managed to collect historical data from CoinMarketCap[23], [24]., including prices, trading volumes, market capitalizations, and other necessary metrics for the bulk of cryptocurrencies. The exploratory data analysis revealed several key findings. The exploratory data analysis revealed some of the significant finding as follows:

a. Cryptocurrency Trends: The pattern of the price trend, the intensity of the market, and the trade volume were observed despite the volatility, again indicating the market's growth [25].
b. Volatility Patterns: Crypto markets[26,27] showed high volatility, which can be necessary for managing risks and diversifying the portfolio.
c. Correlations and Relationships: The study provides a strong and positive correlation between market capitalization and price, with short-term price changes having less influence on the specific variables. This suggests that other external conditions, not just those studied in the paper, drive the market.
d. Comparative Performance: Comparative analysis revealed insights into performance and market activity levels, which gave capital allocation assessments[28].

2) Implications and Discussions: This study's findings have several implications for understanding and navigating the cryptocurrency market. The project's results can be interpreted as having several implications for understanding and successful cryptocurrency investment.

a. Market Maturity: Bitcoin's growth trend is an evident stage in the market's maturation, which is driven by many aspects, such as acceptance, technology, and regulation [29].
b. Risk Management: Increased volatility implies active risk management and broad diversification strategies.
c. Investment Decisions: With a market setting as the performance benchmark in consideration, decision-making is improved.

3) Limitations and Biases: While the project has yielded significant insights, it is essential to acknowledge the limitations and potential biases inherent in the methodology. Although the project's research has produced vital knowledge, it is crucial to recognize the limitations and possible biases that might exist in the formulation:

a. Scraping Approach: It is limited to CoinMarketCap only and, therefore, cannot detect more considerable market tendencies. Along this axis, the mixture of resources could also consist of thematic discussions and social networks as source materials to enrich the analysis.
b. Data Quality: Gathered data may not be consistent, may be missing some parameters, or may have inaccuracies, which can make analysis results unreliable. Errors and bias cannot be prevented even though all the cleaning and data validation are conducted and might go undetected
c. Model Limitations: Forecasting involving predictive models is inherently unstable, affecting projections' long-term accuracy.

While this study has limitations, it nonetheless provides a basis for understanding and predicting cryptocurrency prices. Such knowledge can guide investment strategies and develop models that can adapt to the constantly changing cryptocurrency environment [30].

IV. CONCLUSION

This work shows us the enormous capabilities of combining web scraping tools and data analysis methods into a complex process that allows us to just understand and

forecast the cryptocurrency prices to the smallest detail. Using CoinMarktCap's substantial data resources, the researchers could create a complete data set covering historical prices, trading volumes, market cap, and other essential metrics for many cryptocurrencies.

The knowledge gained from the data exploration and time series methods was related to the cryptocurrency market's behavior. Such dynamics in the growth of Bitcoin's price and market capitalization indicate the worsening adoption rate and technological improvements. Investments in cryptocurrency need risk management and diversification tools that can stand against the risks and volatility, as the analysis shows. Correlation investigations reveal highly correlated links between market capitalization and price and the small correlations of short-term price movements with some other variables, perhaps reflecting the sentiment-driven nature. Comparative analysis guides investment and asset allocation of funds based on the recent reputable cryptocurrency data.

The project's accomplishments, including the conclusions and the implications, have several day-to-day applications. Investors and traders can benefit from the findings to make wise decisions, such as where to put their money, whether to take or hold a risk, or what trading strategies to use on cryptocurrency markets. and the predictive models could also be improved and used in real time for pricing prediction, giving market participants an edge in their ability to foresee and catch the latest price movements and capitalize on them better.

Further research could also aim to integrate data from other sources, namely social media, news articles, and other cryptocurrency-related platforms, as this would enrich the depth and breadth of the analysis being executed. A more thorough analysis of advanced time series forecasting tools, such as the deep learning architecture, will likely yield improved long-term predictive power. Moreover, examining the effect of various external elements, such as legislation modifications, macroeconomic trends, and technological advancements, would provide a broader range of knowledge.

The project has integrated web scraping, data analysis, and predictive modelling to complement the theoretical base of cryptocurrency market analysis, which is its contribution to the field. The framework developed and the lessons learned could provide a useful tool for investors, traders, and researchers working in the fast-changing and complex cryptocurrency industry.

REFERENCES

[1] S. Jani, The Growth of Cryptocurrency in India: Its Challenges & Potential Impacts on Legislation, Faculty of Management Studies, Parul University, Vadodara, India, 2018. doi:10.13140/RG.2.2.14220.36486.

[2] A. Cahyadi, A. I. Owen, F. Ricardo, and A. A. S. Gunawan, "Blockchain Technology behind Cryptocurrency and Bitcoin for Commercial Transactions," in *Proc. 1st Int. Conf. Comput. Sci. Artif. Intell. (ICCSAI)*, Jakarta, Indonesia, 2021, pp. 115–119. doi:10.1109/iccsai53272.2021.9609790.

[3] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten, "SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies," in *Proc. IEEE Symp. Secur. Privacy*, San Jose, CA, USA, 2015, pp. 104–121. doi: 10.1109/SP.2015.14.

[4] A. Syukur and A. Marjuni, "Stock Price Forecasting Using Univariate Singular Spectral Analysis through Hadamard Transform," *Int. J. Intell. Eng. Syst.*, vol. 13, pp. 96–107, 2020. doi:10.22266/ijies2020.0430.10.

[5] N. Bali and A. Singla, "ANN-based wheat crop yield prediction technique for Punjab region," in Emerging Research in Computing, Information, Communication and Applications, N. R. Shetty, L. M. Patnaik, H. C. Nagaraj, P. N. Hamsavath, and N. Nalini, Eds., *Lect. Notes Electr. Eng.*, vol. 790, Springer, Singapore, 2022. doi:10.1007/978-981-16-1342-5_16.

[6] S. Yu, "Comparative analysis of machine learning techniques for cryptocurrency price prediction," *Appl. Comput. Eng.*, vol. 32, pp. 1–12, 2024. doi: 10.54254/2755-2721/32/20230175

[7] M. K. Salman and A. A. Ibrahim, "Price Prediction of Different Cryptocurrencies Using Technical Trade Indicators and Machine Learning," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 928, p. 032007, 2020. doi: 10.1088/1757-899X/928/3/032007

[8] S. Ghimire and H. Selvaraj, "A Survey on Bitcoin Cryptocurrency and its Mining," in *Proc. 26th Int. Conf. Syst. Eng. (ICSEng)*, Sydney, NSW, Australia, 2018, pp. 1–6. doi:10.1109/icseng.2018.8638208.

[9] A. Hellani, A. E. Samhat, M. Chamoun, H. E. Ghor, and A. Serhrouchni, "On Blockchain Technology: Overview of Bitcoin and Future Insights," in *Proc. IEEE Int. Multidiscip. Conf. Eng. Technol. (IMCET)*, Beirut, Lebanon, 2018, pp. 1–8. doi:10.1109/imcet.2018.8603029.

[10] N. S. Zainudin, C.-Y. Ting, K.-C. Khor, K.-H. Ng, G.-K. Tong, and S. N. Kalid, "Clustering Defensive Shariah-compliant Stocks Using Financial Performance as the Indicator," *JOIV: Int. J. Informatics Vis.*, vol. 8, 2024. doi: 10.62527/joiv.8.1.2269.

[11] G. Buana, R. Tyasnurita, N. Puspita, R. Vinarti, and F. Mahananto, "Text-Based Content Analysis on Social Media Using Topic Modelling to Support Digital Marketing," *JOIV: Int. J. Informatics Vis.*, vol. 8, 2024. doi: 10.62527/joiv.8.1.1636.

[12] I. Darmawan, M. Maulana, R. Gunawan, and N. Widiyasono, "Evaluating Web Scraping Performance Using XPath, CSS Selector, Regular Expression, and HTML DOM With Multiprocessing Technical Applications," *JOIV: Int. J. Informatics Vis.*, vol. 6, p. 904, 2022. doi: 10.30630/joiv.6.4.1525.

[13] R. Raju, M. SaiVignesh, and K. I. A. Prasad, "A Study of Current Cryptocurrency Systems," in *Proc. Int. Conf. Comput. Power, Energy, Inf. Commun. (ICCPEIC)*, Chennai, India, 2018, pp. 203–209. doi: 10.1109/iccpeic.2018.8525166.

[14] F. Valencia et al., "Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning," *Entropy*, vol. 21, no. 6, p. 589, 2019. doi: 10.3390/e21060589

[15] S. Oikonomopoulos, K. Tzafilkou, D. Karapiperis, and V. Verykios, "Cryptocurrency Price Prediction using Social Media Sentiment Analysis," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ)*, 2022, pp. 1142–1147. doi: 10.1109/fuzz55282.2022.00143.

[16] G. Gurupradeep, M. Harishvaran, and K. Amsavalli, "Cryptocurrency Price Prediction using Machine Learning," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 12, no. 4, p. 808, 2023. doi:10.17148/ijarcce.2023.124140.

[17] P. Ashok, D. M. Reddy, and A. S. Shaik, "Cryptocurrency price prediction using deep learning algorithms: A comparative study," *Commun. Appl. Nonlinear Anal.*, vol. 32, no. 4s, pp. 537–551, Dec. 2024. doi: 10.52783/cana.v32.2870.

[18] M. Hippke et al., "Wōtan: Comprehensive time-series detrending in Python," *Astron. J.*, vol. 158, p. 143, 2019. doi: 10.3847/1538-3881/ab3984.

[19] B. Amirshahi and S. Lahmiri, "Investigating the effectiveness of Twitter sentiment in cryptocurrency close price prediction by using deep learning," *Expert Syst.*, 2023. doi: 10.1111/exsy.13428.

[20] W. McKinney, Python for Data Analysis, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2017. [Online]. Available: https://www.oreilly.com/library/view/python-for-data/9781491957653/

[21] G. T. Wilson, "Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712," *Journal of Time Series Analysis*, vol. 37, no. 5, pp. 709–711, Mar. 2016, doi:10.1111/jtsa.12194.

[22] A. Shankhdhar, A. K. Singh, S. Naugraiya, and P. K. Saini, "Bitcoin price alert and prediction system using various models," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 1131, p. 012009, 2021, doi:10.1088/1757-899X/1131/1/012009.

[23] S. Bhatt, M. Ghazanfar, and M. Amirhosseini, "Machine learning-based cryptocurrency price prediction using historical data and social media sentiment," in *Proc. 5th Int. Conf. Mach. Learn. Appl. (CMLA)*, Sydney, Australia, Jun. 17–18, 2023, AIRCC Publishing Corporation, doi: 10.5121/csit.2023.131001

[24] T. A. Atwan, Time Series Analysis with Python Cookbook, 2nd ed. Birmingham, UK: Packt Publishing, Apr. 2025. [Online]. Available: https://www.oreilly.com/library/view/time-series-analysis/9781805124283/

[25] N. Suparmanto, A. Asih, A. Sudiarso, and A. Permanasari, "Pricing Decision System for Custom Design Batik," *JOIV: Int. J. Informatics Vis.*, vol. 8, 2024. doi: 10.62527/joiv.8.1.2000.

[26] R. N, S. R. R, V. S. R and K. P. D, "Crypto-Currency Price Prediction using Machine Learning," *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, 2022, pp. 1455-1458, doi:10.1109/ICOEI53556.2022.9776665.

[27] A. Bouteska et al. "Cryptocurrency price forecasting – A comparative analysis of ensemble learning and deep learning methods", *International Review of Financial Analysis*, Volume 92, 2024, 103055, doi:10.1016/j.irfa.2023.103055.

[28] S. Kavithra, V. P. Girishankar, G. Iniyan and B. Logesh, "Cryptocurrency price prediction through integrated forecasting techniques," *2024 3rd International Conference on Artificial Intelligence for Internet of Things (AIIoT)*, Vellore, India, 2024, pp. 1-6, doi:10.1109/AIIoT58432.2024.10574745.

[29] J. Kim, H. Wimmer, H. Liu and S. Kim, "A Streaming Data Collection and Analysis for Cryptocurrency Price Prediction using LSTM," *2021 IEEE/ACIS 6th International Conference on Big Data, Cloud Computing, and Data Science (BCD)*, Zhuhai, China, 2021, pp. 45-52, doi: 10.1109/bcd51206.2021.9581491.

[30] S. Chavan, J. Gundakaram, S. Dyuti Vaishnavi, S. Prasad and K. Deepa, "Real-Time Data Extraction and Prediction of Cryptocurrency," in *IEEE Access*, vol. 12, pp. 186703-186709, 2024, doi: 10.1109/access.2024.3513034.